Cross-Attentive News Integration for Stock Price Forecasting

Nam Huy Nguyen, Thai Binh Nguyen, Van Kien Nguyen, and Anh Cuong Le

Natural Language Processing and Knowledge Discovery Research Group, Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City, Vietnam
522H0131@student.tdtu.edu.vn, 522H0140@student.tdtu.edu.vn, nguyenvankien.st@tdtu.edu.vn, leanhcuong@tdtu.edu.vn

Abstract. Accurately forecasting stock prices remains a challenging task due to the complex interaction between market dynamics and external information such as financial news. Existing multimodal forecasting approaches often process news and market data independently and therefore fail to identify which news events are truly relevant to past price movements. To address this limitation, we propose a cross-attention–based method that explicitly models the relationship between historical price fluctuations and textual news. The approach dynamically focuses on news articles according to their relevance to recent market behavior, allowing the model to highlight price-informative signals while reducing the influence of noisy or unrelated content. Experiments on real-world stock datasets indicate that the proposed cross-attentive integration improves forecasting performance for certain stocks and prediction horizons, leading to notable gains in MAE, RMSE, and directional accuracy in these cases. These results suggest that relevance-aware news integration, guided through cross attention, can enhance stock price prediction under specific conditions, although the improvement is not uniform across all datasets or horizons.

Keywords: Stock Price Forecasting · Financial Time Series Prediction · News Integration · Cross Attention · Financial news

1 Introduction

Forecasting stock prices is a long-standing challenge in financial research due to the highly dynamic and noisy nature of financial markets. Traditional modeling approaches rely primarily on historical price data, yet stock movements are often influenced by external factors such as corporate events, policy announcements, and market sentiment conveyed through news articles. As a result, incorporating textual information has become an increasingly important direction for improving predictive performance. However, a major limitation of existing multimodal methods is that they typically fuse numerical time-series data and news content in a uniform or static manner, without explicitly determining which news items are actually relevant to past price fluctuations. This leads to models that are easily distracted by irrelevant or weakly related information, ultimately reducing their predictive accuracy.

Recent advances in deep learning and attention mechanisms offer new opportunities to better model cross-modal interactions. In particular, cross attention provides a principled way to align two information streams by allowing one modality to selectively attend to the most informative elements of another. In the context of financial forecasting, this mechanism enables historical price movements to guide the selection of relevant news, thereby capturing meaningful price–news relationships that are often overlooked by conventional approaches. Such relevance-driven integration is especially important given the abundance of daily news and the difficulty of manually identifying impactful articles.

Motivated by these observations, this study investigates a cross-attentive approach to integrating financial news with historical market data for stock price forecasting. Our method models how past price dynamics influence the attention weights assigned to textual information, allowing the forecasting model to emphasize price-related news while suppressing irrelevant noise. Experiments conducted on real-world stock datasets demonstrate that this mechanism enhances prediction accuracy across multiple forecasting horizons, outperforming strong baselines based solely on time series or naive multimodal fusion. These results highlight the importance of relevance-aware news integration and suggest that cross-attention–based modeling can provide a more faithful representation of market behavior than traditional multimodal predictors.

2 Related Work

Stock price prediction is essentially a time series forecasting task, as stock prices evolve sequentially and exhibit temporal patterns that can be analyzed to understand market behavior. Traditional and modern approaches, ranging from linear and nonlinear methods to recent machine learning techniques, often rely on identifying recurring short term and long term structures in the data. However, applying these approaches to time series presents several challenges.

In research and practical applications, various methods have been proposed and developed to improve the analysis of economic time series data. Traditional mathematical and statistical approaches such as the ARIMA model [1, 2] and its variants have attracted considerable attention due to their simplicity and high interpretability. However, these models often rely on strict assumptions, such as stationarity, which limits their ability to capture the complex relationships present in time series data.

With the advancement of artificial intelligence, deep learning methods have been increasingly adopted to overcome the limitations of traditional statistical techniques when dealing with nonlinear and nonstationary time series. Deep learning models based on recurrent neural networks (RNN), such as LSTM[3, 4], have been applied to effectively extract long-term temporal dependencies from complex real-world data. In addition, Transformer-based models leverage the Attention mechanism to focus on the most informative parts of the sequence, thereby enhancing the ability to model relationships among variables and values within the time series[5, 6].

Recent studies have begun to incorporate textual data into economic time series analysis to capture additional information that numerical financial data alone cannot fully represent. In financial market forecasting, Farimani et al.[7] introduced a method that extracts features from time varying news sentiment analysis and integrates them with price data to improve the prediction of cryptocurrency prices and foreign exchange rates. Similarly, Reis Filho and colleagues[8] developed a model that combines agricultural commodity price time series with agriculture related news containing selected keywords, thereby enhancing prediction accuracy. In addition, Mou et al.[9] applied this approach to forecast gold prices and exchange rates, where economic information from news served as an important complementary source. Several other studies have also focused on predicting stock price trends by integrating multiple data sources, including financial news, user commentary, and stock price time series[10–13].

In prior studies, textual data and time series data are typically preprocessed and transformed into feature vectors using specialized extraction techniques. Once both types of data have been converted into vector representations, the text vectors are integrated with the time series vectors through various strategies, depending on the characteristics of the task and the architecture of the machine learning model being employed. One of the simplest integration methods is the direct concatenation of text vectors and time series vectors[14, 15].

To enhance the effectiveness of data integration, several studies have employed tensor fusion methods[10, 12], in which feature vectors from textual data and time series data are combined through a tensor product to more fully model the interactions between components of the two modalities.

In addition to the tensor fusion approach, Gating Networks[13] have also been used to integrate textual data and time series data. A Gating Network employs a dynamic mechanism to adjust the contribution weight of each data source. This method helps reduce computational complexity while mitigating overfitting through its ability to selectively filter information.

Another line of research designs fusion layers that employ flexible mechanisms such as self attention[11], cross attention[9], or Graph Neural Networks (GNN) to enhance the ability to model relationships between different data sources.

After integrating the textual and time-series data, the combined dataset is fed into machine learning models for tasks such as forecasting and classification. Depending on the integration process, deep learning models like LSTM[7, 15] can capture temporal patterns, while Transformer-based models[9] learn long-range dependencies. Hybrid architectures are also used to enhance feature learning; for example, CNN–LSTM models[16] extract local semantic patterns and track their temporal dynamics, whereas LSTM–Transformer combinations[17] better model complex contextual relationships. Additionally, graph-based models such as GAT[12] and GCN[11]can learn relational semantics and incorporate temporal information for sequence-based prediction tasks.

Although these approaches show improved analytical performance, time-series analysis methods that integrate data from multiple sources still face several challenges and require further investigation to achieve better results.

3 Background

3.1 Problem Formulation

Let $\{p_t\}_{t=1}^{T}$ denote a univariate stock price series, where p_t represents the closing price at time t. The objective of the forecasting task is to estimate the future price (or return) y_{t+h} at a prediction horizon h by exploiting both historical market information and textual financial news associated with time t.

Market Data Representation

For each time step t, we construct a market feature vector

$$x_t^{(m)} = f_{\text{market}}(p_{t-L+1}, \dots, p_t), \tag{1}$$

where L denotes the look-back window and $f_{\text{market}}(\cdot)$ extracts standard price-based indicators such as returns, volatility, and technical patterns. The resulting market sequence is defined as

$$\mathbf{X}_t^{(m)} = \left[x_{t-L+1}^{(m)}, \dots, x_t^{(m)}\right]. \tag{2}$$

News Data Representation

Let $\mathcal{N}_t = \{n_t^{(1)}, \dots, n_t^{(K_t)}\}$ denote the set of news articles published within a predefined temporal window prior to t. Each article is encoded using a pretrained language model (e.g., BERT) as

$$e_t^{(i)} = f_{\text{text}}(n_t^{(i)}), \quad i = 1, \dots, K_t, \tag{3}$$

yielding a news embedding matrix

$$\mathbf{X}_t^{(n)} = \left[e_t^{(1)}, \dots, e_t^{(K_t)}\right]. \tag{4}$$

Challenge: News–Market Relevance

Although financial news provides abundant contextual information, only a subset of articles is genuinely related to market movements. Many news items may be irrelevant, redundant, or inconsistent with recent price dynamics. Existing multimodal forecasting approaches frequently process market data and textual information independently, making it difficult to determine which news events are informative with respect to historical price behavior.

Cross-Modal Learning Objective

To explicitly capture the dependency between historic price fluctuations and textual content, we aim to learn a mapping

$$F : \left(\mathbf{X}_t^{(m)}, \mathbf{X}_t^{(n)}\right) \rightarrow \hat{y}_{t+h}, \tag{5}$$

such that the model can:

- identify semantic cues in news that correspond to recent market changes;

- assign relevance-aware weights to news articles conditioned on market features;
- suppress noisy or uninformative textual content.

To achieve this, we incorporate a cross-attention mechanism that computes the relevance of each news embedding $e_t^{(i)}$ given market context $x_t^{(m)}$:

$$\alpha_t^{(i)} = \text{Attention}\big(x_t^{(m)}, e_t^{(i)}\big), \tag{6}$$

thereby allowing the model to emphasize news items that are more likely to influence stock movements.

Learning Objective

The model parameters are optimized by minimizing the discrepancy between predicted and actual future prices:

$$\mathcal{L} = \|y_{t+h} - \hat{y}_{t+h}\| \tag{7}$$

The task can thus be formulated as a time-series forecasting problem with explicit cross-modal dependency modeling between historical market patterns and financial news.

3.2 Cross-Attention Mechanism

Attention mechanisms have become central to modern deep learning due to their ability to dynamically allocate representational capacity to the most informative components of the input. While self-attention captures relationships within a single modality, cross-attention extends this principle to interactions *across* different modalities. This property is crucial in multimodal learning tasks, where the objective is to align two heterogeneous sources of information.

Given a query matrix $Q \in \mathbb{R}^{T_q \times d}$ derived from one modality and key–value matrices $(K, V) \in \mathbb{R}^{T_k \times d}$ derived from another, cross-attention computes:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^\top}{\sqrt{d}}\right) V, \tag{8}$$

where $\sqrt{d}$ is a scaling factor ensuring stable gradients.

The core intuition is that each query vector determines how strongly it attends to all key vectors, enabling the model to integrate information selectively. This selective weighting is particularly relevant for stock forecasting with news, as the model must determine: *Which news articles are relevant to the most recent price movements?*

In practice, cross-attention allows the market-derived query representations to emphasize specific semantic patterns in financial news that correlate with market dynamics. For example, if recent price fluctuations exhibit volatility, the cross-attention mechanism can prioritize news related to macroeconomic uncertainty or firm-specific risks. Conversely, irrelevant or diluted textual information receives lower attention weights, reducing noise and improving forecast stability.

Moreover, cross-attention naturally handles variable numbers of news articles, which is a common scenario in real-world financial datasets, because the attention mechanism can aggregate information over sequences of arbitrary length. This capability provides the flexibility required to integrate textual news with continuous time-series data in an effective and consistent manner.

3.3 Long Short-Term Memory Networks

Long Short-Term Memory (LSTM) networks are a specialized form of recurrent neural networks designed to capture long-range dependencies in sequential data. Their architecture mitigates the vanishing-gradient problem that hinders standard RNNs, enabling them to maintain stable representations over extended temporal horizons.

An LSTM cell maintains two forms of state: the hidden state h_t and the cell state c_t. At each time step, the LSTM updates these states using three gates:

- **Forget gate** f_t determines which information from the previous cell state should be discarded.
- **Input gate** i_t regulates how much new information enters the cell state.
- **Output gate** o_t decides which parts of the cell state contribute to the hidden state.

Formally,

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f), \tag{9}$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i), \tag{10}$$

$$\tilde{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c), \tag{11}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t, \tag{12}$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o), \tag{13}$$

$$h_t = o_t \odot \tanh(c_t). \tag{14}$$

LSTMs are widely adopted in financial forecasting tasks due to their ability to capture local temporal dependencies and short-term fluctuations in price series. They are particularly effective at modeling nonlinear temporal patterns—such as momentum effects, reversal intervals, or volatility clusters—often observed in stock markets.

While attention-based Transformers have gained popularity, LSTMs remain competitive for financial time series with limited data or strong short-range dynamics. Their sequential inductive bias ensures that the temporal structure of the price series is preserved without requiring massive training data.

4 Proposed Method

This section presents our hybrid forecasting architecture that integrates numerical stock time-series data with daily financial news. The proposed model consists of four main components: (i) a PhoBERT-based semantic news encoder,

(ii) a positional encoding mechanism that preserves the chronological ordering of multi-day news, (iii) a multi-head cross-attention module that aligns past price movements with temporally indexed news vectors, and (iv) a temporal aggregator and fusion network for multi-step forecasting. The overall architecture of the proposed model is illustrated in Figure 1.

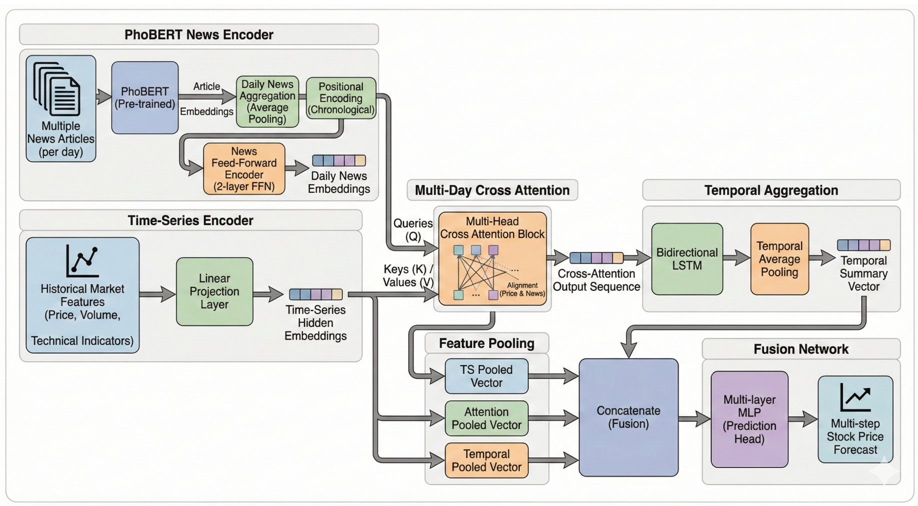

Fig. 1. Architecture of the proposed Cross-Attentive News Integration model for Stock Price Forecasting

4.1 News Representation Using PhoBERT

For each trading day d, we collect all available news articles and encode them using PhoBERT, a pre-trained Vietnamese language model optimized for contextual semantic understanding. Given the embeddings $\mathbf{n}_{d,1}, \mathbf{n}_{d,2}, \ldots, \mathbf{n}_{d,k_d}$ extracted from PhoBERT, where k_d denotes the number of articles on day d, a single day-level representation is obtained by averaging:

$$N_d = \frac{1}{k_d} \sum_{i=1}^{k_d} n_{d,i} \tag{15}$$

This aggregation ensures that each day is associated with one unified semantic vector while retaining the variability introduced by multiple articles.

4.2 Positional Encoding for Chronological Awareness

Because the temporal ordering of news events plays a crucial role in how the market reacts, we integrate sinusoidal positional encoding into the daily news sequence. For a historical window of T days, each day d receives an additive positional vector:

$$\tilde{N}_d = N_d + PE(d) \tag{16}$$

where $\mathrm{PE}(d)$ is the standard sine–cosine positional encoding. This mechanism preserves the chronological structure of news flows and enables the model to differentiate between older and more recent news. The position-aware vectors are subsequently fed into a feed-forward transformation to produce the final enhanced news representations:

$$H_d^{\text{news}} = \mathrm{FFN}(\tilde{N}_d). \tag{17}$$

4.3 Time-Series Encoder

The historical market features, including prices, volumes, and technical indicators, are first projected into the latent space using a linear encoder:

$$H_t^{\text{ts}} = W_{\text{ts}} x_t + b_{\text{ts}} \tag{18}$$

where $\mathbf{x}_t$ denotes the numeric features at day t. The resulting sequence shares the same hidden dimensionality as the news embeddings, which facilitates cross-modal interactions.

4.4 Multi-Day Cross Attention

To explicitly capture the influence of multi-day news on historical price dynamics, we propose a multi-head cross-attention mechanism. Let the time-series encoder produce L query vectors and the news encoder produce M key–value vectors. The cross-attention is defined as:

$$\mathrm{Attention}(Q, K, V) = \mathrm{softmax}\left(\frac{QK^\top}{\sqrt{d}}\right) V. \tag{19}$$

where d is the head dimension.

This component offers three benefits: (1) It aligns price movements with the most relevant news events. (2) It models the cumulative effect of multi-day news flows. (3) It automatically learns the relative importance of news days without manual feature engineering.

The result is a sequence of cross-modal enriched representations reflecting both numerical and event-driven signals.

4.5 Temporal Aggregation

Since the cross-attention module produces a sequence aligned with the time-series window, we employ a bidirectional LSTM to capture higher-order temporal dependencies:

$$Z = \mathrm{BiLSTM}(H_{\mathrm{att}}) \tag{20}$$

where $\mathbf{H}^{att}$ denotes the cross-attention output. A temporal average-pooling is then applied:

$$z_{\mathrm{agg}} = \mathrm{MeanPool}(Z). \tag{21}$$

Similarly, the time-series encoder output and the raw attention output are also pooled:

$$z_{\mathrm{ts}} = \mathrm{MeanPool}(H_{\mathrm{ts}}), \quad z_{\mathrm{att}} = \mathrm{MeanPool}(H_{\mathrm{att}}). \tag{22}$$

4.6 Fusion Network and Forecasting

To fully integrate temporal, numerical, and news-driven information, we concatenate the three pooled vectors:

$$f = [\, z_{\mathrm{ts}} \,;\, z_{\mathrm{att}} \,;\, z_{\mathrm{agg}} \,]. \tag{23}$$

This fused representation is passed through a multi-layer perceptron to yield the final H-step forecast:

$$\hat{y} = \mathrm{MLP}(f). \tag{24}$$

The proposed architecture thus leverages (i) contextual semantic information from PhoBERT, (ii) temporal alignment via positional encoding, and (iii) cross-modal dependence learning through multi-day cross attention. Together, these components enable a more comprehensive modeling of stock price movements influenced by both internal market dynamics and external news events.

5 Experiment

5.1 Dataset

Our experiments are conducted on two complementary datasets: a historical stock price dataset and a financial news corpus. The stock price dataset contains daily trading records of stocks within the VN30 basket from January 1, 2020 to October 20, 2025. The features include fundamental price–volume indicators, technical indicators, and macroeconomic variables. Price-related features were collected directly from Investing.com, while GDP and CPI were obtained from FRED and Investing via API and web scraping, respectively. To avoid data leakage, we intentionally shifted GDP by one year, CPI by one month, and USD/VND by one day relative to the trading date.

The accompanying financial news dataset contains 1,748 articles collected between December 2019 and October 20, 2025. Each entry includes the publication date, article URL, title, category, raw content, and a cleaned content field used for NLP processing. The dataset was adapted and extended from an existing Vietnamese financial news collection originally sourced from Kaggle.

This study employs a temporal split approach to partition the time-series data, preserving the chronological order and preventing data leakage. The dataset is divided into 80% for training and 20% for testing , ensuring that future information is never used during model training. Evaluation is performed using walk-forward validation on the test set, where the model makes iterative one-step-ahead predictions using only historically available information at each time step. This methodology realistically simulates actual trading conditions and provides a robust assessment of the model's predictive capability.

5.2 Experimental Setup

Experimental Environment

All experiments were conducted on the Kaggle computing platform, which provides a standardized and reproducible environment for machine learning research. The experiments were executed using Kaggle's default Python environment with GPU acceleration enabled. The hardware configuration includes an T4 GPU with 15 GB of VRAM, which is sufficient for training deep learning models involving both time-series forecasting and language model–based feature extraction.

Hyperparameter Settings

The model is trained using a batch size of 32, with a constant learning rate of 1×10^{-4} maintained throughout the optimization process. Training is performed for a maximum of 50 epochs. To accommodate the majority of textual inputs, the maximum sequence length is configured to 256 tokens.

5.3 Results

In this section, we present the experimental results obtained from evaluating the proposed forecasting framework with (TS + News) and without news information (TS Only, implemented using an LSTM model). The experiments are conducted on four representative Vietnamese stocks: FPT, ACB, CTG, and DCG. To ensure consistency across all settings, the input window size is fixed at 24, while the forecasting task is performed using four different forecast horizons of 1, 3, 5, and 7 steps ahead. The predictive performance is assessed using three widely adopted evaluation metrics in time series forecasting: MAE, RMSE, and MAPE. These results provide a comprehensive view of how news integration influences forecasting accuracy and reveal the model's behavior across varying prediction horizons.

Table 1. Forecasting performance for FPT stock with and without news

Horizon	**Measure**	**TS + News**	**TS Only**
$h = 1$	MAE	3291.32	5240.48
	RMSE	4097.12	5969.5
	MAPE	2.94	4.78
$h = 3$	MAE	3992.6	7710.9
	RMSE	5004.82	9240.15
	MAPE	3.57	6.79
$h = 5$	MAE	4427.85	9785.85
	RMSE	5635.63	11663.1
	MAPE	3.95	8.6
$h = 7$	MAE	4695.12	10706.85
	RMSE	6048.69	12562.34
	MAPE	4.29	9.43

The results for FPT show that incorporating news consistently improves forecasting accuracy across all horizons. For all metrics (MAE, RMSE, MAPE), the TS + News model performs better than the TS Only baseline, and the advantage remains clear even as the forecasting horizon increases.

Table 2. Forecasting performance for ACB stock with and without news

Horizon	**Measure**	**TS + News**	**TS Only**
$h = 1$	MAE	496.73	259.92
	RMSE	782.63	478.03
	MAPE	2.0	1.05
$h = 3$	MAE	592.8	409.98
	RMSE	886.07	687.56
	MAPE	2.42	1.66
$h = 5$	MAE	704.98	629.64
	RMSE	995.49	1055.84
	MAPE	2.88	2.56
$h = 7$	MAE	720.7	679.47
	RMSE	1063.81	1074.41
	MAPE	2.95	2.78

The results for ACB show that the TS Only model generally achieves better accuracy across most horizons, suggesting that news information does not substantially improve forecasting performance for this stock. This pattern implies that ACB's price movements may be less sensitive to short-term news signals and more strongly driven by historical price behavior.

Table 3. Forecasting performance for CTG stock with and without news

Horizon	Measure	TS + News	TS Only
$h = 1$	MAE	3782.2	2404.53
	RMSE	4517.87	3362.8529
	MAPE	8.6	5.2
$h = 3$	MAE	5602.74	2731.25
	RMSE	6332.78	3956.91
	MAPE	13.0	5.93
$h = 5$	MAE	4110.96	3744.37
	RMSE	4992.62	5147.5
	MAPE	9.34	8.2
$h = 7$	MAE	4267.9147	4431.37
	RMSE	5269.98	5580.86
	MAPE	9.67	9.97

The results for CTG show that incorporating news does not consistently improve forecasting performance across all horizons. For short horizons ($h = 1$ and $h = 3$), the TS Only model achieves lower errors in terms of MAE, RMSE, and MAPE, indicating that short-term price movements are mostly driven by historical prices rather than news. However, for longer horizons ($h = 5$ and $h = 7$), the TS + News model achieves lower RMSE, suggesting that news information can provide additional insights for medium- to long-term forecasting. These results imply that the benefit of relevance-aware news integration is horizon-dependent and may be more pronounced for extended forecasts.

Table 4. Forecasting performance for DCG stock with and without news

Horizon	Measure	TS + News	TS Only
$h = 1$	MAE	3330.16	1261.56
	RMSE	4083.4	1708.08
	MAPE	3.23	1.27
$h = 3$	MAE	3481.59	2645.54
	RMSE	4347.83	3452.57
	MAPE	3.41	2.61
$h = 5$	MAE	2957.26	3914.95
	RMSE	4188.76	5291.94
	MAPE	2.99	3.89
$h = 7$	MAE	3085.28	3626.09
	RMSE	4437.3	5004.8
	MAPE	3.13	3.61

The results for DCG suggest that incorporating news has a limited impact on short-term forecasts but provides noticeable improvements at longer horizons. Specifically, at $h = 1$ and $h = 3$, the TS Only model outperforms the TS + News model in terms of MAE, RMSE, and MAPE, indicating that short-term price movements are primarily driven by historical prices. However, at $h = 5$ and $h = 7$, the TS + News model achieves lower errors across all metrics, particularly RMSE, demonstrating that news information contributes meaningfully to enhancing forecast accuracy over longer horizons.

Overall, the impact of incorporating news on stock forecasting varies across different stocks. For FPT, news consistently improves accuracy across all horizons. In contrast, for ACB, short-term forecasts are mostly driven by historical prices, and news provides little benefit. For CTG and DCG, news information contributes more significantly at longer horizons, while short-term predictions remain dominated by past price movements. These results suggest that the usefulness of news signals depends on both the stock and the forecasting horizon.

6 Conclusion

This study investigates the effectiveness of incorporating news information into stock price forecasting using a combination of time series data and contextual embeddings from PhoBERT. Experimental results on four Vietnamese stocks (FPT, ACB, CTG, and DCG) reveal that the impact of news varies across stocks and forecasting horizons. For FPT, news consistently enhances forecast accuracy across all horizons, while for ACB, short-term forecasts are predominantly driven by historical prices, and news provides limited benefit. For CTG and DCG, news information improves long-term forecasts, particularly for horizons $h \geq 5$, although short-term predictions remain largely determined by past prices.

These findings suggest that integrating textual news signals can be beneficial, especially for longer-term forecasting, but the degree of improvement depends on the stock's sensitivity to news. Future work could explore more sophisticated news aggregation methods, sentiment analysis, or multi-modal approaches combining news, social media, and financial indicators to further enhance predictive performance.

References

1. P. Mondal, L. Shit, and S. Goswami, "Study of effectiveness of time series modeling (arima) in forecasting stock prices," *International Journal of Computer Science, Engineering and Applications*, vol. 4, no. 2, p. 13, 2014.
2. S. Khan, "Arima model for accurate time series stocks forecasting," *International Journal of Advanced Computer Science and Applications*, 2020.
3. H. Widiputra, A. Mailangkay, and E. Gautama, "Multivariate cnn-lstm model for multiple parallel financial time-series prediction," *Complexity*, vol. 2021, no. 1, p. 9903518, 2021.
4. A. Vidal and W. Kristjanpoller, "Gold volatility prediction using a cnn-lstm approach," *Expert Systems with Applications*, vol. 157, p. 113481, 2020.

5. Z. Zeng, R. Kaur, S. Siddagangappa, S. Rahimi, T. Balch, and M. Veloso, "Financial time series forecasting using cnn and transformer," *arXiv preprint arXiv:2304.04912*, 2023.
6. C. Xu, J. Li, B. Feng, and B. Lu, "A financial time-series prediction model based on multiplex attention and linear transformer structure," *Applied Sciences*, vol. 13, no. 8, p. 5175, 2023.
7. S. A. Farimani, M. V. Jahan, A. M. Fard, and S. R. K. Tabbakh, "Investigating the informativeness of technical indicators and news sentiment in financial market price prediction," *Knowledge-Based Systems*, vol. 247, p. 108742, 2022.
8. I. Reis, R. Marcondes, and S. Oliveria, "On the enrichment of time series with textual data for forecasting agricultural commodity prices," p. 101758, 2022.
9. S. Mou, Q. Xue, J. Chen, T. Takiguchi, and Y. Ariki, "Mm-itransformer: A multimodal approach to economic time series forecasting with textual data," *Applied Sciences*, vol. 15, no. 3, p. 1241, 2025.
10. K. Huang, X. Li, N. Xiong, and Y. Yang, "Mf-dat: a stock trend prediction of the double-graph attention network based on multisource information fusion," *Multimedia Systems*, vol. 30, no. 3, p. 136, 2024.
11. Y. Yan, C. Zhang, Y. An, and B. Zhang, "A deep-reinforcement-learning-based multi-source information fusion portfolio management approach via sector rotation," *Electronics*, vol. 14, no. 5, p. 1036, 2025.
12. Q. Zhang, Y. Zhang, F. Bao, Y. Ning, C. Zhang, and P. Liu, "Graph-based stock prediction with multisource information and relational data fusion," *Information Sciences*, vol. 690, p. 121561, 2025.
13. Y. Dong, Z. Wu, and Y. Hao, "Dynamic fusion of multi-source heterogeneous data using moe mechanism for stock prediction," *Applied Intelligence*, vol. 55, no. 6, pp. 1–14, 2025.
14. I. K. Nti, A. F. Adekoya, and B. A. Weyori, "A novel multi-source information-fusion predictive framework based on deep neural networks for accuracy enhancement in stock market prediction," *Journal of Big data*, vol. 8, no. 1, p. 17, 2021.
15. P. Chen, Z. Boukouvalas, and R. Corizzo, "A deep fusion model for stock market prediction with news headlines and time series data," *Neural Computing and Applications*, vol. 36, no. 34, pp. 21 229–21 271, 2024.
16. N. Jing, Z. Wu, and H. Wang, "A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction," *Expert Systems with Applications*, vol. 178, p. 115019, 2021.
17. H. Wang, Z. Xie, D. K. Chiu, and K. K. Ho, "Multimodal market information fusion for stock price trend prediction in the pharmaceutical sector," *Applied Intelligence*, vol. 55, no. 1, p. 77, 2025.